

Atlantic Richfield Company

4 Centerpointe Drive, 2nd Floor, Suite 200
La Palma, CA 906231066
Office: (657) 5294537
Fax: (657) 5294559
E-Mail: Anthony.Brown@bp.com

Anthony R. Brown
Project Manager, Mining

August 28, 2017

Lynda Deschambault
Remedial Project Manager, Superfund Division
U.S. Environmental Protection Agency, Region 9
75 Hawthorne Street, 10th Floor (SFD 7-1)
San Francisco, California 94105

Subject: Response to U.S. EPA's July 27, 2017 Letter "Acceptance of Atlantic Richfield Final Reference Area Work Plan, Revision No. 1 dated April 20, 2017"
Leviathan Mine Site
Alpine County, California

Dear Ms. Deschambault:

Atlantic Richfield Company (Atlantic Richfield) submits this letter in response to the U.S. Environmental Protection Agency's (U.S. EPA's) acceptance of and comments on our Final Reference Area Focused Remedial Investigation Work Plan, Revision No. 1 (Reference Area Work Plan)¹ for the Leviathan Mine Site (site) in Alpine County, California. The Reference Area Work Plan is being implemented in partial fulfillment of the requirements of the Statement of Work attached to the Administrative Order for Remedial Investigation and Feasibility Study [RI/FS], Comprehensive Environmental Response, Compensation, and Liability Act Docket No. 2008-18 issued by the U.S. EPA on June 23, 2008.

In its letter dated July 27, 2017, the U.S. EPA approved the Reference Area Work Plan with the following comments:

- ☐ *EPA understands that the requested reconnaissance mapping planned for the area above the Leviathan Mine Site will be completed this season. Please ensure that the stream reaches identified are sampled; and the data collected is fully considered in the draft Remedial Investigation.*
- ☐ *EPA has reviewed the Reference Area calculation of Reference Threshold Values (RTVs). EPA does not support the use of the upper simultaneous limit (USL) statistic to represent the Reference Threshold Value (RTV). The USL is the upper boundary of the largest value calculated so that all observations in a dataset are less than or equal to the USL, which means RTVs would be higher than the maximum value of a data set. Also, note that The Pro UCL manual referenced is a useful source of information on how to use the software, and technical explanations of statistical tests, but should not be considered EPA policy or guidance. Please utilize the 95th*

¹ Atlantic Richfield, 2017, Final Reference Area Focused Remedial Investigation Work Plan, Leviathan Mine Site, Alpine County, California. Prepared by Amec Foster Wheeler Environment & Infrastructure, Inc., April 20 [Revision No. 1]

percentile or 95% Upper Tolerance Limit (UTL) using the appropriate method based on the data distribution.

Atlantic Richfield disagrees with these comments. For the reasons discussed below, (1) soil samples of stream sediment and floodplain soil in Leviathan and Aspen creeks above the mine are not needed to supplement the existing reference data for these media; and (2) RTVs will be based on 95% USL values calculated using ProUCL software, as stated in each of the three prior versions of the work plan that were reviewed by U.S. EPA.

COLLECTION OF REFERENCE SAMPLES ABOVE LEVIATHAN MINE

As indicated in the Reference Area Work Plan (p. 46), preliminary reconnaissance of physical conditions in Leviathan and Aspen creeks above the site confirmed that these stream reaches were not good analogues for on-property Aspen and Leviathan creeks. As a result, they were not selected as reference riparian areas. The work plan quotes a consistent finding by Dr. David Herbst in his 2002 bioassessment report, which states that “[t]he headwater Leviathan Creek site above the mine is a poor match for habitat setting on impacted sites downstream because the elevation is much higher, the channel is much more narrow, the substrate is dominated by loose gravel and smaller particles, and the temperature is colder.”

Despite this conclusion, the Lahontan Regional Water Quality Control Board requested that further assessment be completed in Aspen and Leviathan creeks above the site. In response, Atlantic Richfield proposed additional reconnaissance-level mapping of floodplain soils and aquatic habitats in those areas. The Reference Area Work Plan indicates that if suitable physical characteristics are observed, plans will be developed for chemical characterization of mapped stream sediment and floodplain soil.

The proposed reconnaissance mapping in upper Aspen and Leviathan creeks is scheduled to occur in late August 2017. However, the recently submitted Stream Sediment Technical Data Summary Report² and Floodplain Soil Technical Data Summary Report³ concluded that no additional reference sediment characterization activities are needed to complete the comparison of stream sediment and floodplain soils collected from reference areas (i.e., Mountaineer and Cottonwood creeks) to stream sediment and floodplain soils in potentially affected reaches of Leviathan and Aspen creeks. Thus, no subsequent chemical characterization is currently anticipated. Sampling plans would only be developed and implemented if the mapping results identify conditions different from those reported by Dr. Herbst in 2002 and observed during preliminary reconnaissance work in 2012 and 2013. Even if samples were collected in late 2017, the results could not be included in draft Site Characterization Report to be submitted in December 2017.

² Atlantic Richfield, 2017a, Stream Sediment Technical Data Summary Report, Leviathan Mine Site, Alpine County, California. Prepared by Amec Foster Wheeler Environment & Infrastructure, Inc., July 14 [Draft]

³ Atlantic Richfield, 2017b, Floodplain Soil Technical Data Summary Report, Leviathan Mine Site, Alpine County, California. Prepared by Amec Foster Wheeler Environment & Infrastructure, Inc., July 14 [Draft]

REFERENCE THRESHOLD VALUES

Atlantic Richfield is stunned to learn that the U.S. EPA does not support the use of the upper simultaneous limit (USL) statistic to represent the Reference Threshold Value (RTV). U.S. EPA reviewed and commented on three prior versions of the Reference Area Work Plan (comments dated July 10, 2015; November 8, 2016; and February 16, 2017) without ever questioning the proposed RTV calculation method. Identifying this as an issue now appears arbitrary, and it raises a host of technical and logistical problems.

Atlantic Richfield expended substantial effort preparing the Reference Area Technical Memorandum.⁴ The reported set of RTVs is a cornerstone of that document. The RTVs were calculated as 95% USLs using the exact methodology presented in each of the work plan drafts and consistent with what we believed were the U.S. EPA's final review comments dated February 16, 2017.

In other letters, the U.S. EPA directed Atlantic Richfield to prepare and submit Technical Data Summary Reports for stream sediment and floodplain soil by July 15, 2017. Each of these documents contained substantial data evaluation that involved the comparison of RTVs, calculated as 95% USLs, to RI/FS metals concentrations in potentially affected reaches of Aspen, Leviathan, and Bryant creeks. Re-running all of those comparisons and re-generating all of the resulting tables, figures, and explanatory text now would require a great deal of work and will unquestionably delay the completion of the Site Characterization Report.

As explained below and in the attached detailed technical response, USLs provide the best indicator of reference conditions for purposes of making the multiple statistical comparisons required for the Site Characterization Report and the RI Report (Singh and Nocerino, 1997⁵). In addition to ProUCL technical guidance, other U.S. EPA guidance supports the use of the USL as a background threshold and specifically, to describe a population "For example, at a polluted site the objective may be to obtain a threshold value estimating the background level contamination prior to any activity that polluted the site. Here, the *upper simultaneous limit*, *USL*, and not the upper confidence limit, *UCL*, for the population mean may be used." (U.S. EPA, 1999⁶). The U.S. EPA is being unduly conservative without providing any technical basis or valid rationale for changing to other statistical metrics as the basis for RTVs. Changing the metric from the USL to the 95th percentile or 95% UTL will result in less meaningful and less defensible RI conclusions.

The 95th percentile and 95% upper tolerance limit (95% UTL) are not appropriate as RTVs because they limit the RTV to a value that describes only a portion of the reference area data (95 percent). Using the 95th percentile or 95% UTL will inappropriately result in sample values being incorrectly identified as exceeding the calculated RTV for the reference population, when

⁴ Atlantic Richfield, 2017, Reference Area Technical Memorandum, Leviathan Mine Site, Alpine County, California. Prepared by Amec Foster Wheeler Environment & Infrastructure, Inc., July 17.

⁵ Singh and Nocerino, 1997, *Chemometrics and Intelligent Laboratory Systems*, v37, 1997. 55–69 *Robust intervals for some environmental applications* Anita Singh, John Nocerino.

⁶ U.S. EPA, 1999, *Robust Statistical Intervals for Performance Evaluations*, Office of Research and Development, November.

they are actually less than the true RTV and should be considered as part of the reference population. In this situation, the misclassification of sample results could lead to incorrect interpretations related to the possible need for remedial actions. Use of the 95th percentile or 95% UTL would inappropriately disregard the purposeful planning that went into the selection of representative reference areas and the exclusion of outliers and evaluation of the data sets that was performed before the RTVs were calculated.

The 95% USL is the most appropriate RTV because it can be used for more than one future comparison, and it provides the most reliable upper bound estimate for reference population concentrations. The RTVs will be compared to an unknown number of data sets representing various study areas, exposure areas, or areas targeted for remedial actions. As a result, U.S. EPA's ProUCL Technical Guide (U.S. EPA, 2015)⁷ suggests that the 95% USL is the most robust and appropriate statistic for these comparisons to control false positive errors. Atlantic Richfield recognizes that the USL must be developed with a representative background data set of one population. As described in the Reference Area Work Plan and the Reference Area Technical Memorandum, the reference area data sets were reviewed for multiple populations and outliers. Outliers were identified and removed using several graphical and statistical methods prior to the calculation of the 95% USLs. Where multiple populations were identified, multiple RTVs were calculated. This controls for excess variability in the reference datasets, adds to the conservatism of the RTVs (controls for "false negatives"), and helps to ensure that we are not overestimating reference metals concentrations. Further detail regarding the basis for the selection of the 95% USLs as RTVs is provided in Attachment A.

Based on the technical justification provided herein and in Attachment A, and to maintain the current schedule to submit the draft Site Characterization Report in December 2017, Atlantic Richfield will continue using 95% USLs as RTVs for comparison to RI/FS metals concentrations in potentially affected media in various study areas associated with Leviathan Mine. We will continue preparing the remaining TDSRs and Draft Site Characterization Report accordingly.

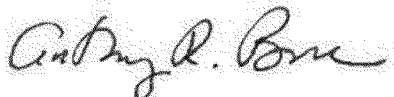
Atlantic Richfield needs immediate confirmation from U.S. EPA of the acceptability of this approach. If U.S. EPA insists on changing course now on how RTVs are to be calculated, it will not be possible to submit the Draft Site Characterization Report by December 31, 2017. In that case, Atlantic Richfield expects that U.S. EPA will modify any applicable Major Submittal deadlines for the RI/FS under Paragraph 50 or pursuant to Section XIX of the UAO, as stated in our June 19, 2017 letter regarding the RI/FS schedule.

⁷ U.S. EPA, 2015, ProUCL Version 5.1.002 Technical Guide, Office of Research and Development, EPA/600/R-07/041, October.

Lynda Deschambault
U.S. Environmental Protection Agency, Region 9
August 28, 2017
Page 5

Please contact me at (657) 529-4537 or anthony.brown@bp.com if we need to discuss this matter further.

Sincerely,

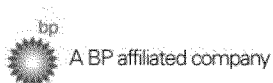


Anthony R. Brown
Project Manager, Mining

Attachment A: Detailed Response Regarding the Use of the Upper Simultaneous Limit as the Reference Threshold Value

cc: Gary Riley, U.S. Environmental Protection Agency, Region 9 – via electronic copy
John Hillenbrand, U.S. Environmental Protection Agency, Region 9 – via electronic copy
Douglas Carey, Lahontan Regional Water Quality Control Board – via electronic copy
Scott Ferguson, Lahontan Regional Water Quality Control Board – via electronic copy
Nathan Block, Esq., BP – via electronic copy
Adam Cohen, Esq., Davis Graham & Stubbs, LLP – via electronic copy
Sandy Riese, EnSci, Inc. – via electronic copy
Marc Lombardi, Amec Foster Wheeler – via electronic copy
Grant Ohland, Ohland HydroGeo, LLC – via electronic copy
Dave McCarthy, Copper Environmental Consulting – via electronic copy
Cory Koger, U.S. Army Corps of Engineers – via electronic copy
Greg Reller, Burleson Consulting – via electronic copy
Ken Maas, U.S. Forest Service, Humboldt-Toiyabe National Forest – via electronic copy and hard copy
Susan Jamerson, Washoe Tribe of California and Nevada – via electronic copy
Neil Mortimer, Washoe Tribe of California and Nevada – via electronic copy
Norman Harry, Washoe Tribe of California and Nevada – via electronic copy and hard copy
Cale Pete, Washoe Tribe of California and Nevada – via electronic copy
Susan Jamerson, Washoe Tribe of California and Nevada – via electronic copy
Fred Kirschner, AESE, Inc. – via electronic copy and hard copy
Ronald Halsey, Atlantic Richfield Company – via electronic copy
Dana Barton, U.S. Environmental Protection Agency, Region 9 – via electronic copy

P:\Project\13000s\13091 Leviathan\4000 Regulatory\4140 RI Work Plans\4 FRIs\8 Reference\170828 RTC EPA 170727 CMTs\170828 RefWP RTC.docx



ED_001709_00000305-00005

ATTACHMENT A

DETAILED RESPONSE REGARDING THE USE OF THE UPPER SIMULTANEOUS LIMIT AS THE REFERENCE THRESHOLD VALUE

The 95% upper simultaneous limit (95% USL) was recommended in the Reference Area Work Plan as the most appropriate reference threshold value (RTV) because it can be used for many future comparisons and provides the best representation of the reference data for these comparisons. The specific technical rationales for selecting the 95% USL and not selecting the other options for reference concentrations proposed by U.S. EPA in their comments (95th percentile and 95% UTL) are provided below.

The 95% USL essentially provides 100 percent coverage of the reference data set with an infinite “n”, for future comparisons (Singh and Nocerino, 1997¹). As described in ProUCL technical guidance (U.S. EPA, 2015²), the use of a USL as an RTV estimate is suggested when a large number of onsite observations (current or future) need to be compared with an RTV. Therefore, if many samples are to be compared and all samples from potentially affected study areas should be correctly identified as less than or greater the reference value, then the 95% USL is the best representation of reference conditions. The RTV was developed from the USL in order to be useful throughout the RI/FS process. It is anticipated for use not only for comparison to existing RI data, but through the Feasibility Study and the selection of cleanup goals. Therefore, the RTVs may need to be compared to an unknown number of data sets representing various potentially affected study areas, exposure areas, or areas potentially targeted for remedial actions. As a result, the 95% USL is the most robust and appropriate statistic for these comparisons to avoid false positive errors. ProUCL technical guidance suggests the use of USLs when many future comparisons are expected (Section 3.1.1. “The use of a USL as a BTV estimate is suggested when a large number of onsite observations (current or future) need to be compared with a BTV).”

Atlantic Richfield recognizes that because of the broad coverage of the 95% USL, it is essential that the USL be developed with a representative background data set. As described in the Reference Area Work Plan and the Reference Area Technical Memorandum, the reference area data sets were reviewed for multiple populations and outliers using several graphical and statistical methods. If outliers were identified, they were removed prior to the calculation of the 95% USLs and in the case of multiple populations, multiple RTVs were calculated. This controls for excess variability in the reference datasets, adds to the conservatism of the RTVs (controls for “false negatives”), and helps to ensure that we are not overestimating reference metals concentrations.

The 95th percentile and 95% upper tolerance limit (95% UTL) are not appropriate because they both limit the RTV to a value that describes only a portion of the reference area data (95 percent). Using either the 95th percentile or 95% UTL will inappropriately result in sample values being incorrectly identified as exceeding the calculated RTV for the reference population, when they are actually less than the true RTV and should be considered as part of the reference population. In this situation, the misclassification of sample results could lead to incorrect interpretations related to the possible need for remedial actions. Further, the number of

¹ Singh and Nocerino, 1997, Chemometrics and Intelligent Laboratory Systems, v37, 1997. 55–69 *Robust intervals for some environmental applications*, Anita Singh, John Nocerino.

² U.S. EPA, 2015, ProUCL Version 5.1.002 Technical Guide, Office of Research and Development, EPA/600/R-07/041, October.

misidentified samples will increase as the number of comparisons increases. This is contrary to the purposeful planning that went into the selection of reference areas, the large numbers of samples collected, and the population evaluation conducted prior to calculation of the 95% USL. In addition, the 95th percentile is not appropriate because, as a percentile, it cannot be used to test other samples and makes no adjustment to account for the confidence in capturing the true background range for future comparisons made to various potentially affected study areas. U.S. EPA's Unified Guidance (Section 6.4.4; U.S. EPA, 2009³) recommends for compliance monitoring that the value be compared to the prediction limit - not the percentile derived from sample data. As described below, the upper simultaneous limit (USL) is a superior statistic to the upper prediction limit (UPL), but was developed in ProUCL after this guidance was published.

We disagree with the use of the 95% UTL as the RTV because of the high number of false positive errors made in future comparisons to datasets from potentially affected study areas. Specifically, UTLs define a value that is expected to be greater than only a set percentage (the coverage) of current and hypothetical future reference data. The default coverage for ProUCL is only 95 percent of the reference data set based on a 95% confidence level (denoted as UTL95-95). This concept is supported by ProUCL technical guidance: "However, when a large number of observations coming from the target population (background, comparable to background) are compared with a UTL95-95, the number of exceedances (not the percentage of exceedances) of UTL95-95 by background observations can be quite large. This implies that a larger number (but not greater than 5 percent) of onsite locations comparable to background may be falsely declared as requiring additional investigation which may not be cost-effective."

Because the 95th percentile and the UTL95-95 are not appropriate for the development of RTVs for this site or any site with a large representative local background or reference data set, ProUCL offers two additional statistical tools for use as background or reference comparison values. These are the UPL and USL. Each is expected to contain a certain percentage of current and future real reference ("background") values with a certain level of confidence. They differ by how many future samples are to be compared.

Because of the expected multiple comparisons of samples from potentially affected study areas to the RTV, we did not recommend the use of the 95% UPL as the RTV because of the probability of making false decisions of reference exceedances during future comparisons (i.e., a false positive error associated with multiple comparisons). U.S. EPA Unified Guidance addresses this issue in Section 6.2.1-The Multiple Comparisons Problem (U.S. EPA, 2009³). Specifically, 95% UPLs are like 95% UTLs but with a 100 percent coverage of background intended for a pre-determined number of future comparisons ("n"). A 95% UPL is essentially a UTL100-95 for just the next "n" samples collected or comparisons made. Therefore, the 95% UPL based upon a background data set is designed to be used to compare "n" future observation(s) with the UPL. ProUCL's default "n" is 1, but it is an option that can be changed by the user. It is expected that the user will know how many comparisons ("n") will be done and enter the appropriate number. Using the UPL for more comparisons than it was calculated for will lead to excessive false exceedances. Unlike the UTL, there are no expected exceedances in the next "n" comparisons, so any exceedance within that "n" is an indication of a non-background sample. However, if more than the assumed number "n" comparisons to the 95% UPL are made, then too many false exceedances are expected to occur. These errors

³ U.S. EPA, 2009, Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Unified Guidance, March 2009, EPA 530/R-09-007.

compound and quickly reach an unacceptable false positive error rate. With a 95% confidence level, each comparison has a 5 percent chance of being a false exceedance. If more than one comparison is done, each comparison has a 5 percent false exceedance chance and after several comparisons, the overall false exceedance chance is much greater than 5 percent. For example, if two subsequent comparisons are performed, the probability of a false exceedance is 9.75 percent, and after only 5 comparisons, it rises to an expected false exceedance rate of greater than 20 percent. So if a 95% UPL for $n=1$ is used for just five comparisons, there is more than a 20 percent chance that at least one comparison will indicate a false exceedance, simply by random chance.